Bacterial metacommunity organization in a highly connected aquatic system

Langenheder, Silke

2017-04

Langenheder , S , Wang , J , Karjalainen , S M , Laamanen , T M , Tolonen , K T , Vilmi , A & Heino , J 2017 , ' Bacterial metacommunity organization in a highly connected aquatic system ' , FEMS Microbiology Ecology , vol. 93 , no. 4 , 225 . https://doi.org/10.1093/femsec/fiw225

http://hdl.handle.net/10138/214437
https://doi.org/10.1093/femsec/fiw225

submittedVersion

Bacterial metacommunity organization in a highly-connected aquatic system

Silke Langenheder[1*], Jianjun Wang[2,3], Satu-Maaria Karjalainen[4], Tiina M. Laamanen[4], Kimmo Tolonen[5], Annika Vilmi[5] and Jani Heino[5]

[1] Department of Ecology and Genetics/Limnology, Uppsala University, Sweden

[2] Department of Geosciences and Geography, University of Helsinki, Helsinki FIN-00014, Finland

[3] State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academic of Sciences, Nanjing 210008, China

[4] Finnish Environment Institute, Freshwater Centre, FI-90570 Oulu, Finland

[5] Finnish Environment Institute, Natural Environment Centre, Biodiversity, FI-90570 Oulu, Finland

for correspondence: silke.langenheder@ebc.uu.se

Note: This is final submitted version.

Abstract

The spatial structure and underlying assembly mechanisms of bacterial communities have been widely studied across aquatic systems, focusing primarily on isolated sites, such as different lakes, ponds and streams. In contrast, biodiversity patterns within large aquatic systems have received less attention. Here, our main aim was therefore to determine the underlying mechanisms for biofilm bacterial assemblages within a large, highly-connected lake system in Northern Finland using associative approaches based on taxonomic and phylogenetic alpha- and beta-diversity and a large number of abiotic and biotic variables. Furthermore, null model approaches were used to quantify the relative importance of different community assembly processes. We found that the spatial variations in bacterial communities within the lake were structured by a combination of different assembly processes, including stochasticity, species sorting and potentially even dispersal limitation. Species sorting by abiotic environmental conditions explained more of the taxonomic and particularly phylogenetic turnover in community composition compared to that by biotic variables. Finally, we observed clear differences in alpha diversity (species richness and phylogenetic diversity), which were to a stronger extent determined by abiotic compared to biotic factors, but also by dispersal effects. In summary, our study shows that the biodiversity of bacterial biofilm communities in a highly-connected lake ecosystem is driven by within-habitat gradients in abiotic conditions as well as by stochastic and deterministic dispersal processes.

Introduction

In recent decades, several related conceptual frameworks have been developed to explain why the composition of ecological communities varies across space (e.g. Leibold et al., 2004; Vellend, 2010). Firstly, spatial variations in community can be due to species sorting (i.e. environmental filtering, habitat filtering or environmental selection) when species are selected by different abiotic and biotic conditions that prevail at different locations. Deviations from pure species sorting can occur if dispersal is limiting, so that species cannot reach locations with suitable conditions (Leibold et al., 2004; Martiny et al., 2006; Nemergut et al., 2013), rendering communities less similar than expected (Stegen et al., 2013). Alternatively, if dispersal rates are so high that they result in mass effects, species are maintained in local communities simply because dispersal rates outpace species sorting processes (Mouquet and Loreau, 2003; Winegardner et al., 2012), leading to at least partly homogenised communities. Finally, local communities can also be stochastically assembled by neutral or drift processes (Hubbell, 2001; Vellend, 2010). This means that a local community is shaped by random differences in birth, death, immigration and emigration among taxa, and hence is simply a random subsample of the regional species pool. Recent literature reviews on bacterial communities have shown that species sorting is the process that often regulates beta-diversity (i.e. differences in community composition between sites) (Hanson et al., 2012; Lindström and Langenheder, 2012; Nemergut et al., 2013). It is nevertheless clear that the other assembly processes can also be important (Hanson et al., 2012; Lindström and Langenheder, 2012; Nemergut et al., 2013; Stegen et al., 2013; Wang et al., 2013). The relative importance of species sorting, compared to other processes, depends on spatial scale at which spatial processes are linked to differences in environmental heterogeneity and dispersal rates within a metacommunity (Martiny et al., 2011; Östman et al., 2012; Wang et al., 2013; Zinger et al., 2014; Comte et al., 2016), which is defined as a set

of local communities that are connected to each other through dispersal (Leibold et al., 2004). For example, Östman et al. (2012) found that bacterioplankton composition can be more strongly associated with stochastic processes in homogeneous compared to heterogeneous environments, because homogeneous environmental conditions do not allow species sorting to occur. Similar conclusions were also made by Wang et al. (2013) based on a comparative survey of several within- and across-habitat studies (including soil, sediments, stream biofilm and lake water). Wang et al. (2013) also presented a conceptual model according to which bacterial communities within habitats should be either assembled by species sorting or stochastic processes, depending on the strength of environmental gradients in the system. In addition, dispersal limitation should be more important in isolated systems, whereas increased connectivity and proximity should lead to increased importance of mass effects in metacommunity organization (Wang et al., 2013; Heino et al., 2015a).

Most studies on community assembly processes in bacterial communities have used statistical approaches, such as distance-decay or variation partitioning methods, where spatial variations in taxonomic community are associated to differences in local environmental conditions and spatial distances between sites, the latter being indicative of dispersal processes (e.g., Martiny et al., 2006; Ramette and Tiedje, 2007). Using such 'associative methods', many studies have shown that both species sorting and spatial processes explain spatial turnover in bacterial communities; however, they often explain only a low fraction of the differences in community composition among sites (Lindström and Langenheder, 2012). It is currently not clear whether this 'unexplained variation' indicates that (a) bacterial communities are to a large extent stochastically assembled, (b) important environmental factors have been missed (Vellend et al., 2014; Heino et al., 2015a), or (c) because they neglect important processes connected to biotic variables, such as the diversity and community composition of other organism groups feeding on bacteria or modifying habitat conditions, which are rarely

96 measured in the field.

97 In parallel, new statistical frameworks based on null model approaches incorporating
98 taxonomic beta-diversity (the total number of species and their relative abundance) and phy-
99 logenetic beta-diversity (species relatedness in a community) have been used to quantify the
100 relative importance of different assembly processes (e.g., Wang et al. 2013, Chase and Myers
101 2011). Stegen et al. (2013; 2015) developed a null model-based analysis framework to disen-
102 tangle the quantitative importance of species sorting, drift, dispersal limitation and mass ef-
103 fects (Stegen et al. 2013). Briefly, the first step is based on the assumption that communities
104 that are phylogenetically clustered have similar traits that have between selected in response
105 to variation in abiotic factors over time (Webb et al., 2002). Hence, if the phylogenetic beta-
106 diversity between a pair of communities deviates significantly from a random null model dis-
107 tribution, it suggests that the communities are assembled by species sorting (that is, 'environ-
108 mental selection' by Stegen et al. (2013)). Consequently, pairs of communities that do not
109 deviate must be assembled by other processes (Stegen et al., 2013). These are then, in the se-
110 cond step, dismantled by determining null model deviations of taxonomic beta-diversity in-
111 stead. Communities that are more similar than expected by chance are assembled by mass
112 effects (that is, 'homogenizing dispersal' by Stegen et al. (2013)), those that are less similar
113 than expected by chance by dispersal limitation, and those that do not deviate from null model
114 prediction by drift (Stegen et al., 2013). There are currently few studies that have used both
115 'associative' and 'null' approaches, and it is therefore difficult to say whether they provide
116 supportive or conflicting results. More generally, it has also been pointed out that there is a
117 need for studies that compare different analytical methods and biodiversity metrics, such as
118 taxonomic and phylogenetic beta-diversity, which provide complementary information on
119 community assembly processes (Jin et al., 2015).
120

Even though bacterial metacommunity organization has been studied widely in aquatic systems in recent years, most studies have focused on patterns of bacterioplankton across isolated water bodies, such as lakes, ponds and rock-pools, or have studied biofilms in stream networks (Wang et al., 2012b; Besemer, 2015; Battin et al., 2016). On the contrary, there are only very few studies on benthic biofilm communities within lakes, where different sampling sites are highly connected (but see Bartrons et al., 2012; Vilmi et al., 2016b). Generally, the biodiversity of biofilms is influenced by local abiotic (i.e. physical and chemical) and biotic conditions, and by immigration of cells from the overlying water column (Besemer, 2015; Battin et al., 2016). The latter process is influenced by passive dispersal processes that transport microorganisms to suitable locations and interactions with the local biofilm community that ultimately determine the colonization success of bacteria from the source community (Besemer, 2015; Battin et al., 2016). Stream biofilm communities are in many circumstances assembled by species sorting processes (Besemer, 2015; Peipoch et al., 2015; Battin et al., 2016), but there are also examples showing that hydrological connectivity and directional flow patterns are important (Liu et al., 2013; Freimann et al., 2015) and may potentially lead to dispersal limitation at larger spatial scales (e.g. Lear et al., 2013). On the contrary, mass effects are of limited importance and cannot 'overrule' species sorting processes even at the level of highly interconnected sites at small spatial scales where the flux of cells between sites is high (Besemer et al., 2009). The degree to which dispersal processes influence the structure of benthic biofilm communities in lakes is currently unclear. In addition, local environmental conditions and dispersal also influence local biodiversity, which has also primarily been studied for bacterial biofilm communities in streams (Besemer, 2015; Battin et al., 2016; Wang et al., 2016). Here, our main aim was therefore to determine the mechanisms that determine alpha- and beta-diversity of biofilm bacterial communities within a large, highly-connected lake ecosystem. We first used associative approaches based on

taxonomic and phylogenetic alpha- and beta-diversity and a large number of abiotic and biotic variables. Second, we apply the null model approach to further quantify the relative importance of different community assembly processes (Stegen et al. (2013)). We hypothesized that within-habitat environmental gradients (both biotic and abiotic) are strong enough to cause differences in alpha- as well as in beta-diversity of biofilm bacterial communities in the lake ecosystem. We further hypothesized that the dispersal of bacteria among sites is high enough to diminish the effects of dispersal limitation on biofilm assembly, but may potentially cause mass effects in some cases.

Material and methods

Study area

The study area, Lake Kitkajärvi, is a large (305 km^2) lake system located in north-eastern Finland where some changes in the water quality and land use have recently been reported (Vilmi et al., 2015). In September 2013, we sampled 36 stony littoral sites for bacteria, algal biomass, diatoms, macroinvertebrates and water (Fig. 1). The 36 sites were as evenly distributed as possible across the perimeter of the whole lake system. The spatial characteristics of the lake system (i.e. the areal extent and high connectivity) enable the organisms to disperse freely among sites.

Bacterial sampling and laboratory procedures

In the field, 10 cobble-sized stones were randomly collected from the water depth of 40 cm at each site. To collect biofilm samples, the surface of each stone was brushed for 20s with a piece of wet foam plastic (4 cm × 4 cm × 4 cm) after which the sample was squeezed into a sampling jar. The samples were immediately stored cold and frozen within the same day.

171 In the laboratory, DNA was extracted from freeze-dried sample material using a
172 PowerSoil DNA Isolation Kit (MOBIO, Carlsbad, USA) and the 16S rRNA gene amplified
173 with the bacterial primers 519f and 926trP1 as described in Vilmi et al (2016b) and sequences
174 on an Ion Torrent PGM™ sequencer (Life Technologies, Gaithersburg, USA).
175 We processed a total number of 404,030 total reads with a mean length of 187 bp
176 mainly using the QIIME pipeline (v1.8) (Caporaso et al., 2010) following previous studies
177 (e.g. Wang et al., 2013). Briefly, the sequences were clustered into OTUs at 97% pairwise
178 identity with the seed-based uclust algorithm (Edgar, 2010). After chimeras were removed via
179 Uchime, representative sequences from each OTU were aligned to the Greengenes imputed
180 core reference alignment V.201308 (DeSantis et al., 2006) using PyNAST (Caporaso et al.,
181 2010). The alignments were then used to construct an approximate maximum-likelihood
182 phylogenetic tree with Jukes-Cantor distance using FastTree (Price et al., 2010) after
183 removing gaps and hypervariable regions using a Lane mask. Taxonomic identity of each
184 representative sequence was determined using the RDP Classifier (Wang et al., 2007) and
185 chloroplast or archaeal sequences were separated out. The lowest sequence depth was 704 and
186 all samples were rarefied to 600 reads for the preparation of the final OTU tables that was
187 used in the alpha- and beta-diversity analyses described below.
188
189 *Biotic and abiotic environmental variables*
190 Biotic variables. At each site, algal biomass was estimated as epilithic phytobenthos
191 chlorophyll *a*, which was measured from the surfaces of 10 stones (collected randomly from
192 40 cm depth) using a BenthoTorch fluorometer (bbe Moldaenke, Cincinnati, USA). Further,
193 diatoms and macroinvertebrates were sampled or surveyed as described in Vilmi et al.
194 (2016b). Diatom samples were brushed from the surfaces of 10 cobble-sized stones from 40
195 cm depth at each site. In the laboratory, permanent slides were made and approximately 500

196 diatom valves were identified to the lowest possible taxonomic level. Macroinvertebrates
197 were sampled using a kick-net with a total kicking effort of 3 min and 6 m at each site and
198 animals were preserved in ethanol. In the laboratory, the animals were sorted and identified to
199 the lowest possible taxonomic level. Further, macroinvertebrates were assigned into different
200 groups based on their feeding habits to separate out biofilm-eating scrapers and their
201 abundance. The following biotic variables were used as predictor variables in the statistical
202 analyses described below: 1) site-specific richness, Shannon's diversity and Pielou's evenness
203 for diatom and macroinvertebrate communities, and biofilm-eating scrapers; 2) the first and
204 second axes of separate non-metric multidimensional scaling analysis (NMDS) for diatom
205 and macroinvertebrate communities, and biofilm-eating scrapers. Finally, we also used the
206 relative abundance of the dominant primary producer *Achnanthidium minutissimum* s.l., as
207 well as the abundance of biofilm-eating scrapers, as biotic predictor variables (see Table S1
208 for a summary).
209 Water chemistry. We performed an extensive sampling campaign for water chemical
210 measurements within two weeks of the sampling of the biotic variables. This time lag was a
211 result of logistic constraints since various chemical variables require immediate analyses in
212 fresh water samples. Samples were taken from a depth of 0.5-1.0 m using a LIMNOS water
213 sampler from a boat at the littoral zone near the bacterial sampling site (i.e. a couple of meters
214 offshore in deeper water to avoid sample contamination by disturbed bottom sediments).
215 Water samples were analyzed within 24 hours of sampling in an accredited laboratory. A total
216 of 35 chemical parameters were analyzed from the site-specific samples (Table S1).
217 Physical characteristics. As physical variables, bottom slope (%) and particle size
218 distribution were measured in the field. Modified Wentworth classes were used to visually
219 assess the coverages of different particle sizes which were mud, fine inorganic sediment (<2
220 mm), gravel (2-16 mm), pebbles (16-64 mm), cobbles (64-256 mm), boulders (256-1024

mm), large boulders (>1024 mm) and bedrock. Subsequently, wind fetch describing the openness of a site was calculated according to Rohweder et al. (2008). For descriptive statistics and abbreviations of the physical variables, see Table S1.

Biodiversity estimators

To determine beta-diversity, we calculated community dissimilarity with and without phylogenetic information. Dissimilarities based on relative abundance data were chosen because they give more weight to dominant OTUs and reduce chance effects that may be involved in the detection of rare OTUs, which may decrease the overall degree of explained variation when presence-absence data are used (Souffreau et al. 2015). Taxonomic turnover was determined using Bray-Curtis dissimilarities based on relative abundances of OTUs between a given pair of samples. To determine phylogenetic turnover, we used the mean nearest taxon distance index (βMNTD) (Fine and Kembel, 2011; Stegen et al., 2012). βMNTD is the mean phylogenetic distance to the closest relative in a paired community for all taxa (Fine and Kembel, 2011) and is sensitive to the changes of lineages close to the phylogenetic tips. Weighted βMNTD based on relative abundance data was calculated according to Stegen et al. (2013).

Bacterial alpha diversity was quantified using species richness and Faith's phylogenetic diversity (PD) (Faith and Baker, 2006).

Statistical analyses

Beta-diversity

To investigate the underlying mechanisms determining beta-diversity in bacterial communities, we used two different approaches. First, we used multiple regression on matrices (MRM) to tease apart the relative importance of sets of variables related to spatial (SPA), abi-

246 otic (ABIO, i.e. chemical and physical) and biotic variables (BIO) on bacterial community
247 similarity. Second, we performed null model analyses according to Stegen et al. (2013) to
248 quantify the relative importance of environmental selection (species sorting), homogenizing
249 dispersal (mass effects), drift and dispersal limitation.
250 *Multiple regression on matrices.* To tease apart the relative importance of the three
251 components (SPA, ABIO and BIO) as well as individual variables on taxonomic and
252 phylogenetic community similarity, we further used the MRM approach (Legendre et al.,
253 1994) with z-score transformed Euclidean distance matrices of the predictor variables
254 (Euclidean distance matrices), as suggested by Martiny et al. (2011). We only included
255 abiotic and biotic variables that were highly correlated (Pearson $r > 0.8$) with community
256 similarities using the "bioenv" function of the "vegan" packages in R (Table S2, (Clarke and
257 Ainsworth, 1993) (Oksanen et al., 2016)). To reduce the effect of spurious relationships
258 between variables, the MRM model was run twice. After the first run, we removed non-
259 significant variables, and we reported the results from the second run only (Martiny et al.,
260 2011). To identify the importance of individual abiotic and biotic variables to the overall
261 correlations, we calculated their partial regression coefficients. Partial regression coefficients
262 provide information about the degree of change in community similarity per standardized unit
263 of similarity for the variable of interest, while all other variables are constant, and thereby
264 identify the variables that make the strongest independent contribution to changes in
265 community composition. The MRM analysis was performed using the R package ecodist
266 v1.2.9 (Goslee and Urban, 2007). Finally, we calculated distance-decay relationships for
267 both taxonomic and phylogenetic beta-diversity to compare their degree of community
268 turnover with increasing spatial distance across sites (Martiny et al. 2006).
269 *Null model analysis.* In the first step, we calculated standardized effect size of βMNTD,
270 which measures (in units of SDs) how much observed βMNTD deviated from the mean of

null distribution (999 null iterations) based on random shuffling of OTU labels across the tips of the phylogeny (Hardy, 2008; Fine and Kembel, 2011; Stegen et al., 2012). This randomization keeps the observed species richness, species occupancy and species turnover constant. We used a significance cut-off of < -2 or > 2, respectively to determine the proportion of community pairs that is phylogenetically more or less similar than expected by chance, respectively. Both cases indicate that environmental selection determines observed compositional differences between samples (Stegen et al., 2013; Dini-Andreote et al., 2015). For all cases where βMNTD did not deviate significantly from the null model distribution (i.e. communities that were not assembled by environmental selection), we calculated the Raup-Crick beta-diversity metric for each pair of local communities after a total of 1,000 iterations (Chase et al., 2011), but based on species relative abundances (RC_{bray}) as in Stegen et al. (2013). Observed RC_{bray} values were compared with those of a random null model distribution according to Chase et al. (2011) and then we followed the procedure described in detail in Stegen et al. (2013) to disentangle the importance of drift, dispersal limitation and mass effects: RC_{bray} values between -0.95 and +0.95 indicate drift, RC_{bray} values > +0.95 indicate that communities are less similar than expected by chance as a result of dispersal limitation, and RC_{bray} values < -0.95 indicate that communities are more similar than expected by chance as a result of mass effects.

Alpha-diversity

Variation in alpha diversity was partitioned between the three components (SPA, ABIO and BIO) using linear models (Borcard et al., 1992; Anderson and Gribble, 1998). By generating models with the three sets of explanatory variables, we estimated the proportions of variation in bacterial diversity explained by the pure effects of SPA, ABIO and BIO, and by the intersections of these three components. For spatial variables, principal coordinates of

296 neighborhood matrices (PCNM; Borcard and Legendre, 2002; Borcard et al., 2004) were used
297 to represent original spatial distance matrices as sets of orthogonal eigenvectors. The first
298 PCNM eigenvector represents the broadest spatial gradient, while each successive eigenvector
299 represents a finer spatial scale. A set of PCNM eigenvectors for each analysis was determined
300 by selecting only positive eigenvectors, which were significant (alpha = 0.05) explanatory
301 variables in a distance-based redundancy analysis DeleteMe model including the eigenvector
302 set and the native spatial distance matrix.

303
304 **Results**

305 *Beta-diversity and community assembly processes*. Significant positive relationships between
306 spatial distance and community dissimilarity were found and the slopes were similar for
307 taxonomic and phylogenetic metrics (Fig. S1). The multiple regression analysis showed that
308 abiotic factors had a relatively stronger effect on both the taxonomic (Bray-Curtis similarities)
309 and phylogenetic (βMNTD) turnover than spatial distance (which had no significant effect) or
310 biotic parameters, where the partial regression coefficients were lower and only marginally
311 significant (Table 1). Moreover, when MRM was run to tease apart the relative importance of
312 individual environmental variables, the partial regression coefficients of algal biomass was
313 significant in the case of taxonomic turnover, whereas no single biotic variable had significant
314 partial regression coefficients in the case of phylogenetic turnover. Among abiotic variables,
315 significant partial regression coefficients were found for NO_x, alkalinity and NH_4 in the case
316 of Bray-Curtis similarities and NO_x in case of βMNTD. Generally, however, the explanatory
317 power of the MRM was low (R^2 values < 0.25 in all cases), so that the largest fraction in
318 differences in community composition remained unexplained.

319 The null model-based approach showed that the majority of pairs of communities were
320 assembled by drift (56% of all pairwise comparisons) whereas 14 % were assembled by

environmental selection, 24 % by dispersal limitation and 6 % by homogenizing dispersal (Fig. 2).

Alpha diversity. At the local scale, species richness and PD at the local scale were significantly correlated to each other ($r^2 = 0.544$, $p < 0.05$, Fig. 3). Generally, a lower fraction of variation in bacterial diversity across sampling sites could be explained by abiotic (7% compared to 18%) and spatial variables (16% compared to 26%) for richness than for PD, respectively. Larger proportion of variation in local diversity were explained by spatial variables than by local environmental conditions, whereas mainly smaller spatial scale variables were significant for species richness and various spatial scale variables for PD. We also found that abiotic variables accounted for significant fractions of variation in both alpha diversity metrics, while the abiotic factors explaining variation in species richness and PD were different (Fig. S2). Richness was negatively correlated to suspended solids and nitrogen, whereas PD was positively correlated to alkalinity, but negatively correlated to aluminium concentrations and Fetch (Fig. S2). Moreover, for PD, we found a significant, albeit minor, effect of biotic variables, as well as considerable co-variation between abiotic, spatial and biotic variables (Fig. 3). Of the biotic variables, PD was positively related to algal biomass and macroinvertebrate richness (Fig. S2).

Discussion

This study shows that within-habitat environmental gradients in one large, highly-connected lake ecosystem were strong enough to cause differences in alpha- and beta-diversity of biofilm bacterial communities. Further, we show that abiotic conditions explained more of the taxonomic and phylogenetic turnover in community composition compared to biotic

346 variables, and that drift, species sorting and dispersal processes contribute to differences in the
347 composition of bacterial biofilm communities between sites.
348
349 Even though it has been shown that within-lake beta-diversity of bacterial
350 communities is lower compared to that between lakes (Yannarell and Triplett, 2004; Wang et
351 al., 2013), it has become clear that there are significant spatial signals for the bacterial
352 communities in sediments (Wang et al., 2013), biofilm (Vilmi et al., 2016b) and lake water
353 (Jones et al., 2012; Lear et al., 2014) within lakes. Data from global surveys in marine
354 systems has also shown that there is stronger community turnover in habitats with stronger
355 environmental gradients (e.g., sediment vs. plankton and coastal vs. open ocean habitats)
356 (Zinger et al., 2011; Zinger et al., 2014). For inland waters, previous studies attributed these
357 spatial patterns in communities to both species sorting and dispersal effects (Wang et al.,
358 2013; Lear et al., 2014; Vilmi et al., 2016b), which was also supported by the results of our
359 study for bacterial biofilms within a lake system. The null model approach showed, moreover,
360 that drift was the predominant assembly process, which fits with conceptual ideas that
361 stochastic assembly should prevail within lakes where environmental gradients should be
362 relatively weak (Wang et al. (2013) and because they are relatively homogenous, well-mixed
363 system where biofilm bacteria are recruited from the water (Besemer, 2015; Battin et al.,
364 2016). At the same time, we might have even underestimated stochastic processes, since they
365 might be masked by indirect species sorting processes by conditions in the water, i.e., species
366 sorting that acted on planktonic bacteria, which then randomly colonized the biofilms. Hence,
367 the variations in biofilm composition reflect the differences in the composition of plankton
368 source communities, which have shown to vary in composition within systems at relatively
369 small spatial scales (Lear et al., 2014). However, as biofilms form at the interface between the
370 substrate (in our case stones) and the water, and are dependent on inorganic nutrients and

371 organic matter from the water to support their growth, it seems unlikely that species sorting
372 effects were indirect. Hence, it has been shown that biofilm assembly is a selective process
373 and not just the results of random dispersal from the surrounding water (Besemer et al., 2012).
374
375 This finding that almost 25 % of all pairwise assemblages are assembled by dispersal
376 limitation is puzzling, in particular because the MRM showed that spatial variables were not
377 significantly related to community similarity. One possible explanation for these deviating
378 results is that the null model analysis overestimates the effect of dispersal limitation. Stegen et
379 al. (2013) define that pairs of communities are assembled by dispersal limitation if RC_{bray} is
380 close to 1, indicating that communities are less similar than expected by chance. This can,
381 however, also be the result of strong biotic forces that create very different communities at
382 adjacent sites (Chase et al., 2011). Here, we found that biotic factors had stronger effects on
383 beta-diversity in the case of taxonomic compared to phylogenetic beta-diversity, and since
384 Stegen et al.'s (2013) approach only used the latter in the identification of species sorting
385 processes, it seems possible that their definition of dispersal limitation to some extent masks
386 the effects of biotic sorting. Another possibility is that the spatial distance matrix that we used
387 in the MRM analyses does not depict the hydrodynamics of our study lake, therefore resulting
388 in non-significant spatial effects. Collectively, the results from different statistical analyses are
389 contradictory, but we cannot currently rule out that dispersal limitation can produce spatial
390 differences in biofilm composition even within a highly-connected lake ecosystem. To fully
391 understand how dispersal influences biofilm community assembly, future studies should
392 therefore utilize new statistical approaches that disentangle effects of directional dispersal
393 through water masses and non-directional processes (e.g. aggregation) at various spatial
394 scales (Bertolo et al., 2012) and integrate direct measurements of water flow rates and
395 directions. In addition to deviating results regarding the importance of dispersal limitation by

396 the MRM and null model approaches, we also found that the results of MRM model differed
397 between taxonomic and phylogenetic beta-diversity. Generally, both biotic and abiotic factors
398 explained less of the similarities in phylogenetic compared to taxonomic beta-diversity. This
399 is, for example, in contrast to a study that compared taxonomic and phylogenetic beta-
400 diversity in a vertical soil gradient, where species sorting was more important in case of the
401 latter (Hu et al., 2015). These deviating results might indicate that the environmental gradients
402 in the highly-connected lake system were not strong enough to select for traits that are
403 phylogenetically conserved, and that community turnover along environmental gradients is
404 therefore better captured by taxonomic diversity metrics.
405
406 Another important finding from our study is that biotic variables were less important
407 than abiotic conditions in structuring biofilm bacterial communities. The most significant
408 biotic factor was algal biomass, which suggests that overall availability of organic substrates
409 released from periphyton influenced bacterial community composition (Wagner et al., 2015;
410 Battin et al., 2016). Algal biomass was significantly correlated to taxonomic but not to
411 phylogenetic turnover, which shows that OTUs irrespective of their phylogenetic affiliation
412 responded to changes in algal biomass. This might reflect the fact that the organic substrates
413 released by periphyton are often highly available and readily used by a wide range of different
414 bacteria (Wagner et al., 2014), and, hence, are not phylogenetically conserved (Martiny et al.,
415 2013). Effects of grazers were generally weak and not significant. For diatom communities in
416 stream biofilms, it has also been shown that abiotic parameters are relatively more important
417 than biotic variables (Göthe et al., 2013). However, the importance of grazers in structuring
418 diatom assemblages was different when the analyses were done for guilds that differed in
419 traits, such as growth form or body size (Göthe et al., 2013). This suggests that the trait
420 composition of the response community also affects the influence of grazers as a structuring

force of the spatial turnover of the grazed communities. Moreover, it might also be possible that the effects of selective grazers, such as protozoa, would have shown a similar picture in our study. Among abiotic variables, nitrite/nitrate concentrations had the strongest independent contribution to the changes in taxonomic and phylogenetic community similarity, confirming previous studies that have shown that nitrogen concentrations are an important structuring factor for biofilm bacterial communities (Kelly et al., 2014; Zeglin, 2015). One methodological caveat of our study was that there was a gap between the sampling for biotic variables (including biofilm composition) and abiotic variables, which were sampled two weeks later. The fact that abiotic factors nevertheless could explain differences in biofilm composition suggests that the time of the sampling was not a major problem. However, it is possible that we actually underestimated the contribution of abiotic factors, which would further strengthen our conclusion because they were more important in structuring biofilm communities compared to biotic factors.

Interestingly, alpha diversity were also affected by similar environmental factors which were most important for determining beta-diversity. However, different factors correlated with the variation in taxonomic richness and PD across sampling sites, and the relationship between species richness and PD, $r^2 = 0.544$, was relatively weak compared to other studies (e.g. Wang et al., 2012a). This is in congruence with unimodal productivity-diversity relationships, which have also been found in bacteria (Song et al., 2016). PD, on the contrary, was positively related to Mg concentration, algal biomass and macroinvertebrate richness. This shows that PD is higher in more benign alkaline environments and that interactions with other trophic groups may promote the co-existence of a phylogenetically diverse bacterial biofilm community as well. Moreover, high algal biomass and macroinvertebrate richness may reflect thicker and more mature biofilm that provide more

446 physical niches and resources. This idea is supported by the finding that fetch, reflecting
447 effects of physical disturbances by wave action, was negatively related to PD. Moreover, Al
448 concentrations were negatively related to PD, which many generally reflect negative effects of
449 metal pollution on biofilm richness (Bier et al., 2015). Taken together, the results suggests that
450 adaptation to environmental pollution and resistance towards strong physical disturbances
451 require traits that might be phylogenetically conserved (Martiny et al., 2015) and therefore
452 decreases PD. Opposite to what we found for beta-diversity, spatial variables had a stronger
453 effect than environmental variables on differences in alpha diversity between sites. This may
454 indicate that the total number of species found in a locality can be influenced by dispersal
455 (including rates and pathways) that transports taxa to a location (Besemer et al., 2013; Zha et
456 al., 2016), but that most taxa remain inactive and contribute primarily to seed-bank at local
457 sites and only occasionally to the active community (Shade et al., 2014). We are, however,
458 aware of the fact that the sequencing depth in our study might have been too low to obtain
459 robust estimates of alpha-diversity (Lundin et al., 2012), so these results need to be
460 interpreted with care.
461
462 To summarize, there is now an increasing number of studies that show clear turnover
463 of bacterial communities at relatively small spatial scales within aquatic ecosystems. Such
464 turnover can be attributed to a combination of different assembly processes, including
465 stochasticity, species sorting and potentially even dispersal limitation. More studies are,
466 however, still needed to integrate the importance of different community assembly processes
467 at different spatial scales depending on connectivity patterns and dispersal pathways between
468 and within different types of inland water bodies.
469
470 **Funding**

This study was funded by the European Regional Development Fund through the Kitka-MuHa project. Annika Vilmi was supported by the Maj and Tor Nessling Foundation, Jianjun Wang was supported by the Emil Aaltonen Foundation and Silke Langenheder by funding from the Swedish Research Council.

Acknowledgements

We thank Yrjö Kivinen, Susan Kunnas, Minna Kuoppala, Sari Latvala, Marja Lindholm, Osmo Murtovaara and Juha Riihimäki for sampling and data collection as well as Marko Suokas from Biocenter Oulu Sequencing Center.

References

Anderson, M.J., and Gribble, N.A. (1998) Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. *Aust J Ecol* **23**: 158-167.

Bartrons, M., Catalan, J., and Casamayor, E.O. (2012) High Bacterial Diversity in Epilithic Biofilms of Oligotrophic Mountain Lakes. *Microbial Ecology* **64**: 860-869.

Battin, T.J., Besemer, K., Bengtsson, M.M., Romani, A.M., and Packmann, A.I. (2016) The ecology and biogeochemistry of stream biofilms. *Nature Reviews Microbiology* **14**: 251-263.

Bertolo, A., Blanchet, F.G., Magnan, P., Brodeur, P., Mingelbier, M., and Legendre, P. (2012) Inferring Processes from Spatial Patterns: The Role of Directional and Non-Directional Forces in Shaping Fish Larvae Distribution in a Freshwater Lake System. *PLOS ONE* **7**: e50239.

Besemer, K. (2015) Biodiversity, community structure and function of biofilms in stream ecosystems. *Research in Microbiology* **166**: 774-781.

Besemer, K., Singer, G., Hodl, I., and Battin, T.J. (2009) Bacterial Community Composition of Stream Biofilms in Spatially Variable-Flow Environments. *Applied and Environmental Microbiology* **75**: 7189-7195.

Besemer, K., Singer, G., Quince, C., Bertuzzo, E., Sloan, W., and Battin, T.J. (2013) Headwaters are critical reservoirs of microbial diversity for fluvial networks. *P Roy Soc B-Biol Sci* **280**.

Besemer, K., Peter, H., Logue, J.B., Langenheder, S., Lindstrom, E.S., Tranvik, L.J., and Battin, T.J. (2012) Unraveling assembly of stream biofilm communities. *ISME J* **6**: 1459-1468.

Bier, R.L., Voss, K.A., and Bernhardt, E.S. (2015) Bacterial community responses to a gradient of alkaline mountaintop mine drainage in Central Appalachian streams. *Isme Journal* **9**: 1378-1390.

Borcard, D., and Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol Model* **153**: 51-68.

Borcard, D., Legrende, P., and Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology* **73**: 1045-1055.

Borcard, D., Legendre, P., Avois-Jacquet, C., and Tuomisto, H. (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology* **85**: 1826-1832.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K. et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335-336.

Chase, J.M., Kraft, N.J.B., Smith, K.G., Vellend, M., and Inouye, B.D. (2011) Using null models to disentangle variation in community dissimilarity from variation in alpha-diversity. *Ecosphere* **2**: Article 24.

Clarke, K.R., and Ainsworth, M. (1993) A Method of Linking Multivariate Community Structure to Environmental Variables. *Mar Ecol Prog Ser* **92**: 205-219.

Comte, J., Lovejoy, C., Crevecoeur, S., and Vincent, W.F. (2016) Co-occurrence patterns in aquatic bacterial communities across changing permafrost landscapes. *Biogeosciences* **13**: 175-190.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K. et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb* **72**: 5069-5072.

Dini-Andreote, F., Stegen, J.C., van Elsas, J.D., and Salles, J.F. (2015) Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences of the United States of America* **112**: E1326-E1332.

533 Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST.
534 *Bioinformatics* **26**: 2460-2461.
535 Faith, D.P., and Baker, A.M. (2006) Phylogenetic diversity (PD) and biodiversity
536 conservation: some bioinformatics challenges. *Evolutionary Bioinformatics* **2**: 121-128.
537 Fine, P.V.A., and Kembel, S.W. (2011) Phylogenetic community structure and phylogenetic
538 turnover across space and edaphic gradients in western Amazonian tree communities.
539 *Ecography* **34**: 552-565.
540 Freimann, R., Burgmann, H., Findlay, S.E.G., and Robinson, C.T. (2015) Hydrologic linkages
541 drive spatial structuring of bacterial assemblages and functioning in alpine floodplains.
542 *Frontiers in Microbiology* **6**.
543 Goslee, S.C., and Urban, D.L. (2007) The ecodist package for dissimilarity-based analysis of
544 ecological data. *J Stat Softw* **22**: 1-19.
545 Göthe, E., Angeler, D.G., Gottschalk, S., Lofgren, S., and Sandin, L. (2013) The Influence of
546 Environmental, Biotic and Spatial Factors on Diatom Metacommunity Structure in
547 Swedish Headwater Streams. *PLOS ONE* **8**.
548 Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., and Martiny, J.B.H. (2012) Beyond
549 biogeographic patters: processes shaping the microbial landscape. *Nat Rev Microbiol*
550 **10**: 497-506.
551 Hardy, O.J. (2008) Testing the spatial phylogenetic structure of local communities: statistical
552 performances of different null models and test statistics on a locally neutral community.
553 *J Ecol* **96**: 914-926.
554 Heino, J., Melo, A.S., Siqueira, T., Soininen, J., Valanko, S., and Bini, L.M. (2015a)
555 Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns,
556 processes and prospects. *Freshw Biol* **60**: 845-869.
557 Hu, W.G., Zhang, Q., Tian, T., Li, D.Y., Cheng, G., Mu, J. et al. (2015) Relative Roles of
558 Deterministic and Stochastic Processes in Driving the Vertical Distribution of Bacterial
559 Communities in a Permafrost Core from the Qinghai-Tibet Plateau, China. *PLOS ONE*
560 **10**: e0145747.
561 Hubbell, S.P. (2001) *A unified neutral theory of biodiversity and biogeography*. Princeton, NJ:
562 Princeton University Press.
563 Jin, L.N.S., Cadotte, M.W., and Fortin, M.J. (2015) Phylogenetic turnover patterns consistent
564 with niche conservatism in montane plant species. *Journal of Ecology* **103**: 742-749.
565 Jones, S.E., Cadkin, T.A., Newton, R.J., and McMahon, K.D. (2012) Spatial and temporal
566 scales of aquatic bacterial beta diversity. *Front Microbiol* **3**: 318.

Kelly, J.J., Minalt, N., Culotti, A., Pryor, M., and Packman, A. (2014) Temporal Variations in the Abundance and Composition of Biofilm Communities Colonizing Drinking Water Distribution Pipes. *PLOS ONE* **9**: e98542.

Lear, G., Bellamy, J., Case, B.S., Lee, J.E., and Buckley, H.L. (2014) Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. *ISME J* **8**: 1715-1726.

Lear, G., Washington, V., Neale, M., Case, B., Buckley, H., and Lewis, G. (2013) The biogeography of stream bacteria. *Global Ecology and Biogeography* **22**: 544-554.

Legendre, P., Lapointe, F.J., and Casgrain, P. (1994) Modeling Brain Evolution from Behavior - a Permutational Regression Approach. *Evolution* **48**: 1487-1499.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F. et al. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* **7**: 601-613.

Lindström, E.S., and Langenheder, S. (2012) Minireview: Local and regional factors influencing bacterial community assembly. *Environ Microbiol Rep* **4**: 1-9.

Liu, J., Soininen, J., Han, B.P., and Declerck, S.A.J. (2013) Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *Journal of Biogeography* **40**: 2238-2248.

Lundin, D., Severin, I., Logue, J.B., Ostman, O., Andersson, A.F., and Lindstrom, E.S. (2012) Which sequencing depth is sufficient to describe patterns in bacterial alpha- and beta-diversity? *Environ Microbiol Rep* **4**: 367-372.

Martiny, A.C., Treseder, K., and Pusch, G. (2013) Phylogenetic conservatism of functional traits in microorganisms. *ISME Journal* **7**: 830-838.

Martiny, J.B.H., Jones, S.E., Lennon, J.T., and Martiny, A.C. (2015) Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**.

Martiny, J.B.H., Eisen, J.A., Penn, K., Allison, S.D., and Horner-Devine, M.C. (2011) Drivers of bacterial beta-diversity depend on spatial scale. *P Natl Acad Sci USA* **108**: 7850-7854.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L. et al. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Rev Microbiol* **4**: 102-112.

Mouquet, N., and Loreau, M. (2003) Community patterns in source-sink metacommunities. *Am Nat* **162**: 544-557.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F. et al. (2013) Patterns and Processes of Microbial Community Assembly. *Microbiology and Molecular Biology Reviews* **77**: 342-356.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B. et al. (2016) vegan: Community Ecology Package. R package version 2.3-3. https://cran.r-project.org/package=vegan.

Östman, Ö., Drakare, S., Kritzberg, E.S., Langenheder, S., Logue, J.B., and Lindstrom, E.S. (2012) Importance of space and the local environment for linking local and regional abundances of microbes. *Aquat Microb Ecol* **67**: 35-U158.

Peipoch, M., Jones, R., and Valett, H.M. (2015) Spatial Patterns in Biofilm Diversity across Hierarchical Levels of River-Floodplain Landscapes. *Plos One* **10**.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE* **5**: e9490.

Ramette, A., and Tiedje, J.M. (2007) Biogeography: An emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microb Ecol* **53**: 197-207.

Rohweder, J., Rogala, J.T., Johnson, B.L., Anderson, D., Clark, S., Chamberlin, F., and Runyon, K. (2008) Application of wind fetch and wave models for habitat rehabilitation and enhancement projects. *US Geological Survey*: Open-File Report 2008–1200, 2043 p.

Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N., and Gilbert, J.A. (2014) Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. *Mbio* **5**.

Song, W., Kim, M., Tripathi, B.M., Kim, H., and Adams, J.M. (2016) Predictable communities of soil bacteria in relation to nutrient concentration and successional stage in a laboratory culture experiment. *Environmental Microbiology* **18**: 1740-1753.

Stegen, J.C., Lin, J., Konopka, A.E., and Fredrickson, J.K. (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* **6**: 1653-1664.

Stegen, J.C., Lin, X.J., Fredrickson, J.K., and Konopka, A.E. (2015) Estimating and mapping ecological processes influencing microbial community assembly. *Frontiers in Microbiology* **6**.

Stegen, J.C., Lin, X.J., Fredrickson, J.K., Chen, X.Y., Kennedy, D.W., Murray, C.J. et al. (2013) Quantifying community assembly processes and identifying features that impose them. *ISME J* **7**: 2069-2079.

Vellend, M. (2010) Conceptual Synthesis in Community Ecology. *Q Rev Biol* **85**: 183-206.

Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleynhans, E.J. et al. (2014) Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos* **123**: 1420-1430.

Vilmi, A., Karjalainen, S.M., Landeiro, V.L., and Heino, J. (2015) Freshwater diatoms as environmental indicators: evaluating the effects of eutrophication using species morphology and biological indices. *Environmental Monitoring and Assessment* **187**.

Vilmi, A., Karjalainen, S.M., Nokela, T., Tolonen, K., and Heino, J. (2016) Unravelling the drivers of aquatic communities using disparate organismal groups and different taxonomic levels. *Ecol Indic* **60**: 108-118.

Wagner, K., Besemer, K., Burns, N.R., Battin, T.J., and Bengtsson, M.M. (2015) Light availability affects stream biofilm bacterial community composition and function, but not diversity. *Envir Microbiol* **17**: 5036-5047.

Wagner, K., Bengtsson, M.M., Besemer, K., Sieczko, A., Burns, N.R., Herberg, E.R., and Battin, T.J. (2014) Functional and Structural Responses of Hyporheic Biofilms to Varying Sources of Dissolved Organic Matter. *Appl Environ Microbiol* **80**: 6004-6012.

Wang, J., Meier, S., Soininen, J., Casamayor, E.O., Feiyan, P., Tang, X. et al. (2016) Regional and global elevational patterns of microbial species richness and evenness. *Ecography* **DOI: 10.1111/ecog.02216**.

Wang, J.J., Soininen, J., He, J.Z., and Shen, J. (2012a) Phylogenetic clustering increases with elevation for microbes. *Env Microbiol Rep* **4**: 217-226.

Wang, J.J., Soininen, J., Zhang, Y., Wang, B.X., Yang, X.D., and Shen, J. (2012b) Patterns of elevational beta diversity in micro- and macroorganisms. *Global Ecology and Biogeography* **21**: 743-750.

Wang, J.J., Shen, J., Wu, Y.C., Tu, C., Soininen, J., Stegen, J.C. et al. (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J* **7**: 1310-1321.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microb* **73**: 5261-5267.

Webb, C.O., Ackerly, D.D., McPeek, M.A., and Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics* **33**: 475-505.

Winegardner, A.K., Jones, B.K., Ng, I.S.Y., Siqueira, T., and Cottenie, K. (2012) The terminology of metacommunity ecology. *Trends in Ecology & Evolution* **27**: 253-254.

667 Yannarell, A.C., and Triplett, E.W. (2004) Within- and between-lake variability in the
668 composition of bacterioplankton communities: investigations using multiple spacial
669 scales. *Appl Environ Microbiol* **70**: 214-223.
670 Zeglin, L.H. (2015) Stream microbial diversity in response to environmental changes: review
671 and synthesis of existing research. *Front Microbiol* **6**: 454.
672 Zha, Y.H., Berga, M., Comte, J., and Langenheder, S. (2016) Effects of Dispersal and Initial
673 Diversity on the Composition and Functional Performance of Bacterial Communities.
674 *Plos One* **11**.
675 Zinger, L., Boetius, A., and Ramette, A. (2014) Bacterial taxa-area and distance-decay
676 relationships in marine environments. *Mol Ecol* **23**: 954-964.
677 Zinger, L., Amaral-Zettler, L.A., Fuhrman, J.A., Horner-Devine, M.C., Huse, S.M., Welch,
678 D.B.M. et al. (2011) Global Patterns of Bacterial Beta-Diversity in Seafloor and
679 Seawater Ecosystems. *PLOS ONE* **6**.
680
681
682
683
684

685 **Table 1.** Multiple regression analysis on distance matrices for taxonomic (Bray- Curtis) and
686 phylogenetic beta diversity metrics (βMNTD). Overall R^2 values as well as partial regression
687 coefficients for sets of abiotic (ABIO), biotic (BIO) and spatial (SPA) variables selected by
688 the bioenv analysis, as well as individual ABIO and BIO variables are shown. Note that only
689 variables with significant partial regression coefficient of at least one of the beta-diversity
690 metrics ($p < 0.05$) are listed. Partial regression coefficients and p-values for all variables are
691 shown in Table S4.

	Bray- Curtis	**p-value**	**βMNTD**	**p-value**
Sets of variables				
R^2	***0.217***	***0.000***	***0.189***	***0.011***
ABIO	0.025	0.001	0.012	0.021
BIO	0.011	0.044	0.005	0.047
SPA	0.002	0.699	-0.001	0.643
Individual variables				
R^2	***0.241***	***0.000***	***0.217***	***0.039***
Alkalinity	0.026	0.000		
NH_4^+	-0.028	0.010		
NO_x-N	0.043	0.000	0.012	0.028
Particle mean size				
Algal biomass	0.10	0.018		

692

693

694 **Figure legends**
695
696 **Figure 1.** The study area with 36 littoral sampling sites.

697

698 **Figure 2**. Proportions of community pairs assembled by drift, species sorting (Selection),

699 dispersal limitation (Disp. Lim) and mass effects or homogenizing dispersal (Hom. Disp.).

700

701 **Figure 3.** (A) Relationship between species richness and phylogenetic diversity. (B) and (C)

702 Variation in bacterial species richness (B) and phylogenetic diversity (C) related to spatial

703 distance (SPA), abiotic (ABIO), and biological variables (BIO). Inorganic SS: inorganic sus-

704 pended solids (µg L^{-1}), NOx: NO2+NO3-N (µg L^{-1}), Mg: Mg concentration (mg L^{-1}), Al:

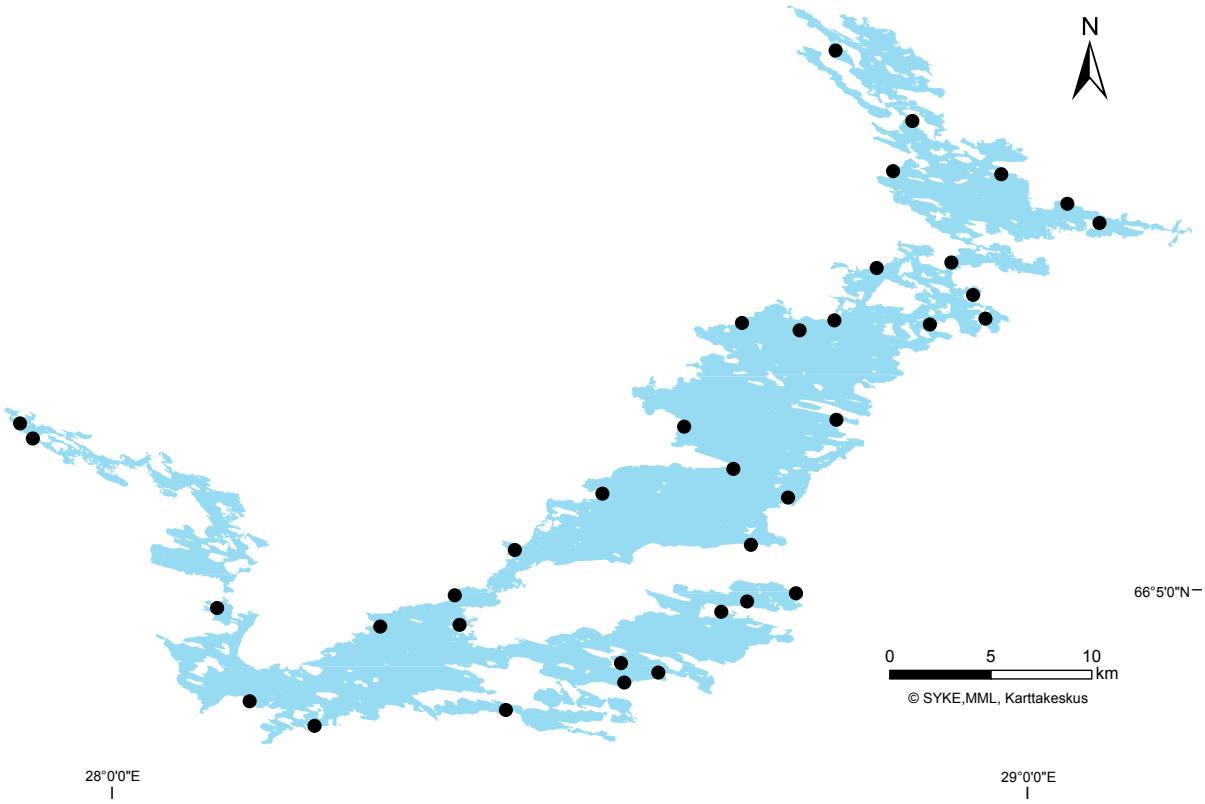
705 Aluminum concentration (mg L^{-1}), Macro.comm.richness: Species richness of macroinverte-

706 brates. PCNM eigenvectors were determined as described in the material and methods.

707

708

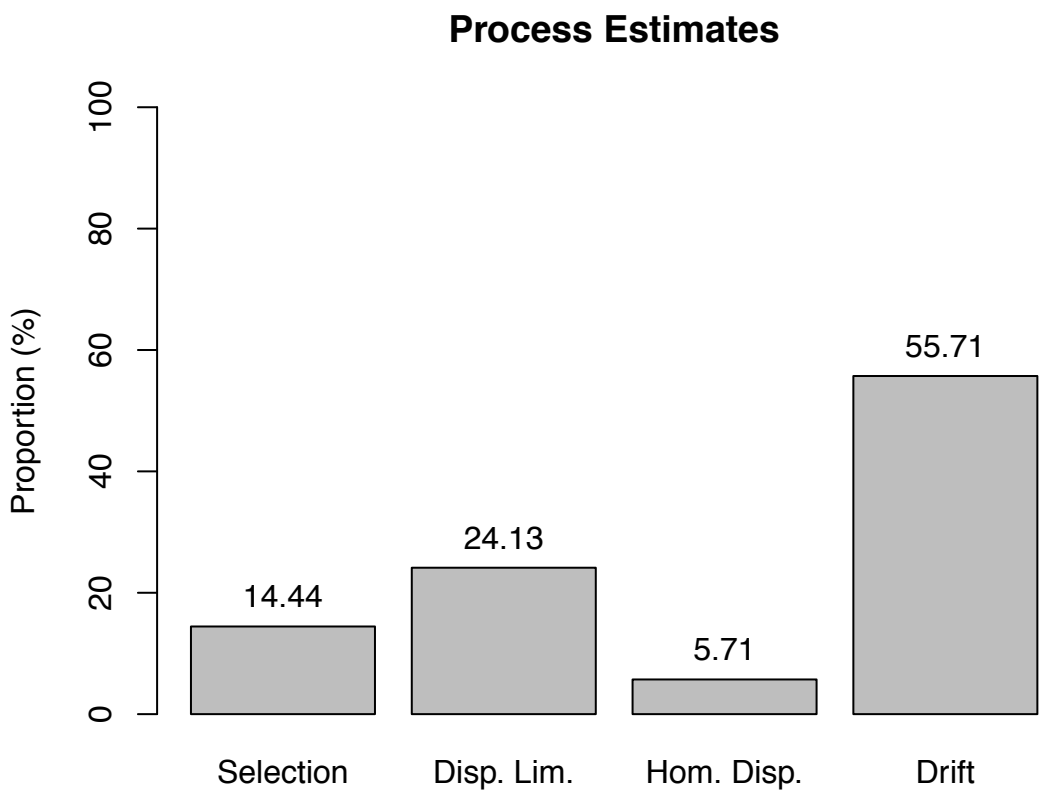
709 **Figure 1**



710

711

712 **Figure 2**



713

714

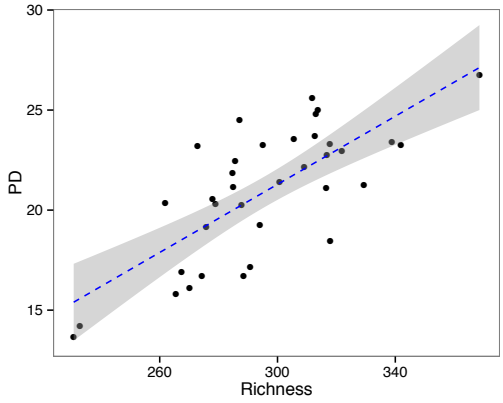
715

716

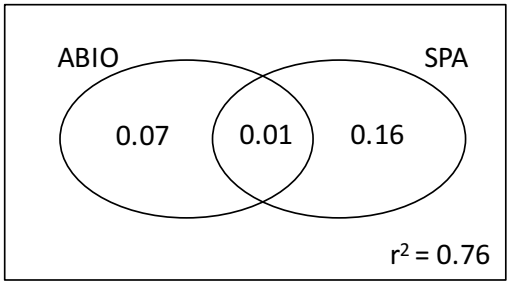
717 **Figure 3**

718

A. Relationship between Richness and PD

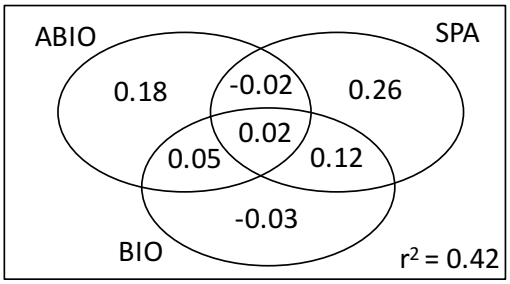


B. Variation partitioning: Richness



ABIO: Inorganic SS, NO_x; p = 0.049
BIO: ns
SPA: PCNM 16, 14, 4; p = 0.018

C. Variation partitioning: PD



ABIO: Mg, Fetch, Al; p = 0.004
BIO: Algal Biomass, macro.comm.richness; p = 0.025
SPA: PCNM 4, 16, 14, 3, 1, 2; p = 0.004

719

720